



武汉大学学报(信息科学版)

Geomatics and Information Science of Wuhan University

ISSN 1671-8860, CN 42-1676/TN

《武汉大学学报(信息科学版)》网络首发论文

题目: 点云数据直线检测及其在人工林树木计数中的应用
作者: 方浩, 李红军
DOI: 10.13203/j.whugis2021040710.13203/j.whugis20210407
收稿日期: 2021-07-29
网络首发日期: 2022-07-18
引用格式: 方浩, 李红军. 点云数据直线检测及其在人工林树木计数中的应用[J/OL]. 武汉大学学报(信息科学版).
<https://doi.org/10.13203/j.whugis2021040710.13203/j.whugis20210407>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

DOI:10.13203/j.whugis20210407

引用格式:

方浩, 李红军. 点云数据直线检测及其在人工林树木计数中的应用[J]. 武汉大学学报·信息科学版, 2022, DOI: 10.13203/j.whugis20210407 (FANG Hao, LI Hongjun. Counting of Plantation Trees Based on Line Detection of Point Cloud Data[J]. *Geomatics and Information Science of Wuhan University*, 2022, DOI: 10.13203/j.whugis20210407)

点云数据直线检测及其在人工林树木计数中的应用

方浩¹, 李红军¹

¹北京林业大学理学院, 北京, 100083

摘要: 基于激光点云数据进行人工林调查时, 由于激光扫描时树木的遮挡与自遮挡、树木被砍伐等原因造成扫描的点云数据有缺失, 遗漏树木的位置判断不准确, 森林调查结果误差大。解决这一问题的关键是实现缺失树木的填补。本文定义了离散点集共线度的概念, 构建了一个基于点集共线度最大化模型并结合直线检测进行缺失数据填补的方法。模拟数据实验结果: 本文方法的平均准确率为 97.28%。人工林数据实验结果: 本文方法检测到 9 棵缺失树的位置, 共线度由 0.2193 增大为 0.2705。实验结果表明, 本文方法不仅可以实现缺失位置的最优推断, 加强填补后数据的共线关系, 也可应用于人工林的缺失树木计数。

关键词: 离散点集; 直线检测; 缺失数据填补; 树木计数; 共线度

中图分类号: P237; TP751.1 文献标志码: A

1. 引言

森林调查是森林经理和森林监控的常规工作^[1]。而树木的计数则是森林调查的工作内容之一, 有助于监控树木的资源数量和生长质量, 直接影响森林材积计算和碳储量的估计。近几十年来, 随着信息技术的发展, 采用遥感影像和激光扫描进行森林调查已

经成为研究的热点话题之一^[2-5]。其中, 激光扫描数据因具有准确的三维信息, 在树木位置信息和几何形状信息提取方面具有优势。此外, 采用森林点云数据进行三维重建是计算机图形领域的一个热点研究话题^[6-7]。重建的森林三维几何模型可以用于三维(3D)动画、电影、游戏、园林景观设计、虚拟旅

¹ 收稿日期: 2021-07-29

项目资助: 本研究受北京林业大学 2020 年研究生课程建设项目资助(HXKC2005); 国家自然科学基金(61571046)。

第一作者: 方浩, 硕士研究生, 主要从事点云分析与处理方法的研究。2505119182@qq.com

通讯作者: 李红军, 博士, 教授。lihongjun69@bjfu.edu.cn

游等。

无论是森林树木计数还是重建三维数字森林,都需要完整的森林点云数据。但是,由于激光扫描时树木的遮挡和自遮挡、树木被砍伐等原因造成扫描的点云数据有缺失,不仅树木的计数不准确,而且重建的森林也不完整。解决这一问题的途径之一是针对不完整点云数据进行填补修复。

对于不完整点云,进行缺失分析和缺失数据修复都是很有挑战性的工作^[8]。首先,森林点云的数据量很大,单站激光扫描采集的点云包含数百万甚至上千万的离散点,因此数据存储量大,计算量也很大;其次,不容易确定缺失树木位置和大小,即使是人工林,树木大致成行成列的网格形式栽植,但是,树木之间的间距也不是严格相等,点云数据中的行列直线检测并不是一件容易的事情;第三,由于野外施工要求并不严格或者由于地理条件限制,往往树木并不完全对齐,这会导致树木共线检测的误差较大。

点云数据的采集用到的激光扫描仪价格昂贵且在森林作业成本高,如果因数据采集不完整而直接删除数据,会造成资源的浪费,而且若数据的缺失不是完全随机的,直接删除存在缺失的数据会导致后期分析结果存在较大的偏差。因此,处理缺失数据的

首选方案是将其进行填补。常用的填补算法有最大期望算法^[9]、热卡填补法^[10]、回归填补法^[11]、聚类填充^[12]、机器学习算法^[13-14]等。前四类填补算法进行数据填补时需要利用缺失样本的部分已知信息,否则填补效果不佳或无法进行填补。机器学习方法^[15]虽然可以根据数据特征之间的关系对各种数据进行填补,但往往需要大量的同类样本进行模型训练。

分析并利用点云数据内蕴的几何结构信息进行数据缺失分析和填补显然是有广泛的适用性^[8],而直线或者线段的识别是各种结构分析的基础。点云数据中的直线检测是近些年热点研究方向之一^[16]。比较有代表性的直线检测方法可以分为四类:最小二乘法^[17]、随机抽样一致性(Random Sample Consensus, RANSAC)^[18]、Hough 变换^[19-20]和其他方法,比如 LSD(Line Segment Detector)直线检测算法^[21-22]。最小二乘法检测直线精度高,但要求处理数据集中噪声点较少。RANSAC 算法速度快,能找出噪声较大的数据中的直线,但一般只检测出单一直线且随机性较高。Hough 变换^[23]具有对局部缺损的不敏感优良特性,但对参数选择的依赖性较高。LSD 算法^[24]检测速度快,错检率低,但对于误差较大的点云数据检测效果差,

需要后处理。

为了提高点云数据直线检测的准确性以及解决应用共线特征进行缺失数据填补的问题,本文提出一种离散点集共线度的概念及其计算方法,并基于共线度构建一种缺失数据填补算法,该算法以共线度最大化为目标建立优化模型,实现在考虑误差容忍度条件下缺失树木位置自动推断。

2. 共线度和直线检测方法

2.1. 共线度的定义

给定平面上离散点集 $\Omega = \{p_i(x_i, y_i), i = 1, 2, \dots, n\}$ 和有限直线集 $L = \{l_1, l_2, \dots, l_m\}$, n 为集合 Ω 中点的个数, 即 $|\Omega| = n$, m 为直线数, 即 $|L| = m$ 。

定义 1 经过点 p_i 的所有直线 l 所经过的离散点集 Ω 中的点的个数占总点数 n 的比值称为点 p_i 关于点集 Ω 和直线集 L 的共线度, 简称为点 p_i 的共线度, 记为 $deg(p_i)$, 即

$$deg(p_i) = \frac{1}{n} \sum_{j=1}^n \mathcal{K}_L(p_i, p_j) \quad (1)$$

其中 $\mathcal{K}_L(p_i, p_j)$ 为定义在直线集 L 上的示性函数, 即

$$\mathcal{K}_L(p_i, p_j) = \begin{cases} 1, \exists l \in L, \text{使得 } p_i \text{ 和 } p_j \text{ 同在 } l \text{ 上} \\ 0, \text{其他情况} \end{cases} \quad (2)$$

从定义可知, 点的共线度的大小与直线

集有关。

定义 2 点集 Ω 中, 所有点的共线度的均值称为点集 Ω 关于直线集 L 的共线度, 简称为点集 Ω 的共线度, 记为 $deg(\Omega)$ 。

$$deg(\Omega) = \frac{1}{n} \sum_{i=1}^n deg(p_i) \quad (3)$$

2.2. 直线检测算法

对于没有给定直线集的平面离散点集 Ω , 则首先需要进行直线集检测, 才能计算点的共线度。而对于离散点集, 因研究问题的需求不同, 需要指定检测的直线至少需要经过的点的个数。

定义 3 离散点集 Ω 中, 由至少经过 k 个点确定的直线称为 k -直线。当 $k = 2$ 时, 任意 2 个点确定一条直线, 称之为平凡直线。

实际工程研究中, 主要研究非平凡直线的检测。

定义 4 设点集 P 是点集 Ω 的一个非空子集, 若 P 中所有点在同一条直线上, 则称 P 为 Ω 的一个共线子集。

如何从离散点集中求出所有 k -直线 ($k > 2$) 是一个具有挑战性的计算几何问题。这是一个 NP-hard 问题。此外, 工程实践中, 测量的数据往往不是严格共线, 而是误差在允许的范围, 比如不超过 ϵ 时, 都算共线。为了避免严格的直线检测在工程实践应用中造成

遗漏和冗余,本文提出一个带共线误差容忍度 ε 的离散点集 Ω 的 **k -直线检测算法(k -Line Detection Algorithm, 简称为 k -LDA)**, 该算法的主要步骤如下:

第 1 步: 输入数据和参数。即输入点集 Ω , 指定直线至少经过的点数 k 和共线误差容忍度 ε 。

第 2 步: 计算共线子集。对 $\forall p_i, p_j \in \Omega$ 且 $p_i \neq p_j$, 则 p_i 和 p_j 确定的直线记为直线 l_{ij} , 集合 $\{p_i, p_j\}$ 记为 P_{ij} 。再遍历 Ω 中的所有点, 按照规则: $\forall p_c \in \Omega, c \neq i$ 且 $c \neq j$, 计算离散点 p_c 到直线 l_{ij} 的距离 d_c , 若 $d_c < \varepsilon$, 则 $p_c \in P_{ij}$; 否则 $p_c \notin P_{ij}$ 。

第 3 步: 合并重复共线子集。对共线子集 P_{ij}, P_{nm} , 若 $|P_{ij} \cap P_{nm}| \geq 2$, 则判定它们重合, 将两个子集合并为新子集。重复执行第 3 步, 直到所有子集之间都不重合。

第 4 步: 拟合直线集。对共线子集 P_{ij} , 若 $|P_{ij}| \geq k$, 利用最小二乘法对 P_{ij} 进行直线拟合, 所有拟合直线构成直线集 L ; 若 $|P_{ij}| < k$, 则剔除该共线子集。

从算法流程可以看出, 我们的算法时间复杂度为 $\tau \cdot O(n^4)$, 其中 τ 为第 3 步循环次数, 最大为 $\frac{n(n-1)}{2}$ 。这是一个非高效算法, 当点数比较少时是可用的, 如何提高时间效率是未

来一个有价值的研究方向。

3. 基于共线度最大化的缺失数据填补算法

对于多重共线的数据(比如人工林的树木位置), 即每一个数据点都是 2 条或者 2 条以上直线的交点的数据, 若有数据缺失, 则可以利用共线度进行缺失数据填补。根据共线度和 **k -LDA** 算法, 本文构造了一种针对多重共线数据的缺失数据填补算法, 称为**多重共线位置缺失数据填补 (Filling in multiple collinear position missing data, FMCP) 算法**, 其主要步骤如下:

第 1 步: 计算直线集 L 。对于二维离散点集 Ω , 其本身并不包含与之对应的直线集。因此, 在计算其共线度之前需根据 **k -LDA** 算法求出直线集 L ;

第 2 步: 指定有效范围 Γ 。设定数据的有效范围, 不考虑有效范围外的数据, 使缺失数据分析和填补范围有明确的边界;

第 3 步: 确定备选集。确保数据的多重共线性不变, 规定缺失的离散点是 k -直线的交点, 把有效范围 Γ 内 k -直线的交点集合称为备选集, 记为 \bar{Q} ;

第 4 步: 确定缺失点集 Q 。确保填补后数据之间的最短距离不变, 从 \bar{Q} 中找出使得点

集 $deg(\Omega \cup Q)$ 最大的缺失数据点集 Q 。

实际上,第4步缺失点集 Q 的确定是一个0-1规划问题。设 $\bar{Q} = (q_1, q_2, \dots, q_{|\bar{Q}|})$,决策变量为 $\omega_i, i = 1, 2, \dots, |\bar{Q}|$.其中

$$\omega_i = \begin{cases} 1, & q_i \in Q \\ 0, & q_i \notin Q \end{cases} \quad (4)$$

根据 $deg(\Omega \cup Q)$ 应该尽可能大,建立优化模型(5):

$$\begin{cases} \max \frac{1}{|\Omega|^2} \left(\sum_{i=1}^{|\Omega|} \sum_{j=1}^{|\Omega|} \mathcal{K}_L(p_i, p_j) + \sum_{i=1}^{|\Omega|} \sum_{j=1}^{|\bar{Q}|} (2\mathcal{K}_L(p_i, q_j) \cdot \omega_j) \right. \\ \left. + \sum_{i=1}^{|\bar{Q}|} \sum_{j=1}^{|\bar{Q}|} (\mathcal{K}_L(q_i, q_j) \cdot \omega_i \cdot \omega_j) \right) \\ s.t. \quad \omega_i + \omega_j \leq 1, \text{若 } d(q_i, q_j) \leq \rho, i, j = 1, 2, \dots, |\bar{Q}|, \text{且 } i \neq j \\ \omega_i = 0, \text{若 } \exists p \in \Omega, \text{使得 } d(p, q_i) \leq \rho, i = 1, 2, \dots, |\bar{Q}| \\ \omega_i = 0 \text{ 或 } 1, i = 1, 2, \dots, |\bar{Q}| \end{cases} \quad (5)$$

$d(\cdot)$ 表示欧氏距离, ρ 为距离限制参数。此处的共线度计算与3.1的定义略有差异,这是因为点集 $\Omega \cup Q$ 中点的个数不确定,为了方便模型的求解,使用 $|\Omega|$ 代替 $|\Omega \cup Q|$ 。

模型(5)是一个非线性0-1规划问题,求解困难,时间成本较高,下面通过约简目标函数的方法求出其近似最优解。

在计算各点共线度时,不考虑缺失离散点之间共线关系,可以实现对模型(5)目标函数的非线性项进行简化,形成模型(6)如下

$$\begin{cases} \max \frac{1}{|\Omega|^2} \left(\sum_{i=1}^{|\Omega|} \sum_{j=1}^{|\Omega|} \mathcal{K}_L(p_i, p_j) + \sum_{i=1}^{|\Omega|} \sum_{j=1}^{|\bar{Q}|} (2\mathcal{K}_L(p_i, q_j) \cdot \omega_j) \right) \\ s.t. \quad \omega_i + \omega_j \leq 1, \text{若 } d(q_i, q_j) \leq \rho, i, j = 1, 2, \dots, |\bar{Q}|, \text{且 } i \neq j \\ \omega_i = 0, \text{若 } \exists p \in \Omega, \text{使得 } d(p, q_i) \leq \rho, i = 1, 2, \dots, |\bar{Q}| \\ \omega_i = 0 \text{ 或 } 1, i = 1, 2, \dots, |\bar{Q}| \end{cases} \quad (6)$$

这是一个线性0-1规划问题,其求解方法简单^[25]。对模型(6)求解得最优解集 Φ 。该最优解集中,若最优解是唯一的,则该解就作为模型(5)的近似最优解 $\omega^* = (\omega_1^*, \omega_2^*, \dots, \omega_{|\bar{Q}|}^*)$;若最优解不唯一,则通过模型(5)目标函数中的二次项部分,进行优化,建立模型(7)

$$\begin{cases} \max \frac{1}{|\Omega|^2} \left(\sum_{i=1}^{|\bar{Q}|} \sum_{j=1}^{|\bar{Q}|} (\mathcal{K}_L(q_i, q_j) \cdot \omega_i \cdot \omega_j) \right) \\ s.t. \quad (\omega_1, \omega_2, \dots, \omega_{|\bar{Q}|}) \text{ 为模型(6)的解} \end{cases} \quad (7)$$

求解模型(7),得 $\omega^* = (\omega_1^*, \omega_2^*, \dots, \omega_{|\bar{Q}|}^*)$,实现从模型(6)的解 Φ 中选出模型(5)的近似最优解 ω^* 。

于是,可以根据 $Q = \omega^* \cdot \bar{Q}$ 计算出缺失离散点集。

4. 缺失数据填补实验

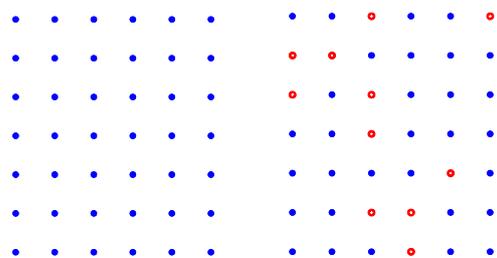
为了验证基于共线度最大化的缺失数据填补算法(FMCP)的有效性和优势,我们用模拟数据和实际数据进行了算法实验,前者有利于量化的分析。使用Hough变换替代FMCP中的k-LDA算法作为对比算法,简称为Hough+算法。由于Hough算法是图像直线检测算法,按照规则:一平方单位面积平分10×10个像素和规则:若像素点内存在离散点,则其值为1,否则为0将二维离散点集变换为二值化图像,图像的大小由离

散点集的范围确定。Hough 变换的参数空间细分参数 $[d\rho, d\theta]=[1,1]$ ，单位分别为像素和度；邻域大小为 $[3,15]$ ，单位分别为像素和度。寻找峰值时的阈值为直线通过最少点数 k 和查找的峰值的最大数量 T 。

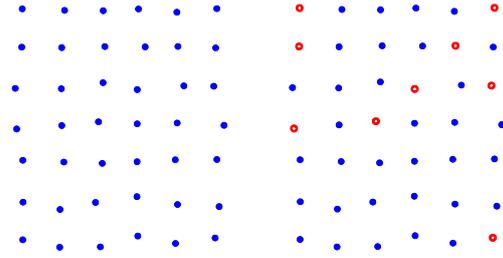
4.1. 模拟数据实验

取每个离散点到其他点的最小距离的平均值为点间距 R ，参数 $\rho = 0.95R$ ，误差限 $\varepsilon = 0.05 R$ ，区域 Γ 为矩形域 $\{(x, y) | x_{min} \leq x \leq x_{max}, y_{min} \leq y \leq y_{max}\}$ ，其中 $x_{min}, x_{max}, y_{min}, y_{max}$ 为已有离散点 p 的坐标最值，直线通过最少点数 $k = 4$ (k 值取太大会加强直线检测条件，可能出现漏检情况；而 k 值取太小会减弱直线检测条件，可能出现错检情况)。

生成的模拟数据是间隔为1的标准网格离散点，其规格为 $N_r \times N_c$ ，如图 1(a)。从中随机选择一定比例 σ ($0 \leq \sigma < 1$)的数据作为缺失点，如图 1(b)中红色圆环点，其余数据作为已有离散点，如图 1(b)中蓝色点。



(a) 网格点 (b) 红圈为随机去除点



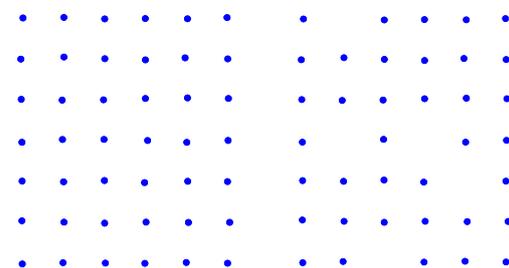
(c) 带噪声网格点 (d) 带噪声含缺失点集

图 1 模拟数据

Fig.1 Simulated Data

考虑到实际中离散点带有一定偏差，在生成网格离散点时，加入随机噪声参数 α 。在标准网格点 α -邻域内随机生成的离散点，如图 1(c)。同样地，随机从中选择一定比例 σ 的数据作为待补充离散点，如图 1(d)中红色圆环点，其余数据作为已有离散点，如图 1(d)中蓝色点。

当 $(N_r, N_c, \sigma, \alpha) = (6, 7, 0.1, 0.05)$ 时，算法过程如图 2 所示。红色圆环点为检测到的缺失数据位置。从图 2(d)中可以看出，本文缺失数据填补算法将缺失数据全部填补，且填补位置与原始点集相近。



(a) 生成网格点 (b) 随机去除点

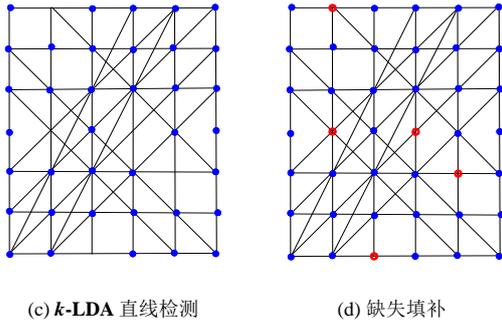


图 2 基于 FMCP 算法的模拟数据的缺失填补

Fig.2 Missing Recovery of Simulated Data with FMCP Algorithm

为了定量分析实验结果的好坏，本文采用完整率 (Completeness Rate, CR) 和准确率 (Accuracy Rate, AR) 两种定量指标进行评价。

完整率的定义

$$CR = \frac{ps}{pc} \times 100\% \quad (8)$$

式中: ps 为填补点属于缺失点的个数, pc 为真实缺失点的个数。

准确率的定义

$$AR = \frac{ps}{ps + pe} \times 100\% \quad (9)$$

式中: pe 为填补点不属于缺失点的个数。

对不同的 $(N_r, N_c, \sigma, \alpha)$ 分别进行 50 次重复实验, 平均实验指标结果列在表 1 和表 2 中, 其中 Hough+ 算法中参数 $T = 20N_r$ 。

表 1 不同算法的完整率 (%) 比较

Tab.1 Comparison of the Completeness Rate (%) of Different Algorithms

σ	α	$N_r \times N_c$												
		6×7		7×8		8×9		9×10		10×11		11×12		
		FMCP	Hough+	FMCP	Hough+	FMCP	Hough+	FMCP	Hough+	FMCP	Hough+	FMCP	Hough+	
0.1	0	97.53	91.86	100	99.29	100	98.30	100	100	100	100	100	99.54	100
	0.05	52.83	92.86	86.38	96.49	98.80	99.46	99.34	99.78	99.14	100	99.85	100	
	0.1	16.74	22.08	39.84	60.90	67.91	72.53	79.75	83.49	92.31	87.59	92.55	90.98	
	0.2	0.97	8.57	4.48	17.89	21.69	44.01	39.82	64.69	64.90	74.00	78.57	74.61	
0.2	0	86.97	74.16	96.36	89.74	99.03	95.52	100	99.46	100	100	100	99.85	
	0.05	29.10	64.25	67.39	89.49	86.44	97.69	94.56	99.56	98.99	99.38	99.40	99.76	
	0.1	8.17	14.03	19.96	32.96	40.38	52.51	64.47	69.29	80.65	78.08	89.41	83.55	
	0.2	0.69	2.84	1.23	12.07	10.65	26.87	21.15	46.79	43.93	63.35	60.47	65.89	
0.3	0	56.58	37.24	85.86	66.67	95.58	90.43	98.68	97.09	99.63	98.28	99.95	99.95	
	0.05	16.69	44.43	32.56	62.93	66.28	89.55	81.65	93.45	89.47	98.03	96.83	99.38	
	0.1	1.95	5.95	10.11	15.29	22.75	32.41	40.09	46.21	58.44	56.93	74.75	70.50	
	0.2	0.32	0.46	1.34	4.20	4.90	12.72	12.03	25.00	21.84	42.14	39.65	54.11	

表 2 不同算法的准确率 (%) 比较

Tab.2 Comparison of Accuracy Rate (%) of Different Algorithms

σ	α	$N_r \times N_c$											
		6×7		7×8		8×9		9×10		10×11		11×12	
		FMCP	Hough+	FMCP	Hough+	FMCP	Hough+	FMCP	Hough+	FMCP	Hough+	FMCP	Hough+
0.1	0	100.0	99.51	100	100	100	99.43	100	100	100	100	100	100
	0.05	99.12	99.49	99.62	99.64	99.39	100	100	99.78	100	100	100	100
	0.1	97.44	100	100	100	99.22	98.88	98.97	98.38	99.60	98.62	99.83	99.03
	0.2	100	94.74	92.86	91.67	92.31	76.70	89.80	74.18	88.97	77.08	86.84	73.14
0.2	0	100	96.88	100.00	98.10	100.00	98.94	100.00	99.78	100	100	100.	100
	0.05	100	98.47	98.85	100.00	99.33	99.69	99.66	99.89	99.81	99.91	100	99.92
	0.1	100	98.21	98.21	96.69	98.68	97.66	99.30	96.94	98.47	96.80	98.64	97.93
	0.2	100	85.71	84.62	87.01	92.50	85.12	86.04	85.11	89.60	81.15	86.18	75.41
0.3	0	100	95.02	100.00	96.71	99.91	98.76	100	99.45	100	99.88	100	100
	0.05	100	98.92	98.94	98.91	98.30	99.79	99.37	99.85	99.67	99.76	99.58	99.95
	0.1	100	100	95.70	94.12	97.17	93.62	95.68	95.98	97.03	93.47	97.80	95.45
	0.2	100	100	100	78.26	92.73	84.57	85.64	81.60	86.30	82.52	86.33	81.12

从表 1 可以看出,在完整率这个指标上, Hough+算法在噪声水平较高时优于本文算法,而本文算法在噪声水平低时有优势。从表 2 可以看出,在准确率这一个指标上,本文算法在噪声水平高时明显优于 Hough+算法;本文算法的平均准确率为 97.28%,高于 Hough+算法的平均准确率(94.77%)。

4.2. 人工林数据实验

数据采集自河北省威县国家重点杨树良种基地的一块三倍体毛白杨人工林样地(图

3(a))。该样地经过单站激光扫描得到该人工林三维点云数据,该数据包含 27904956 采样点,分布在一个长 175 米,宽 125 米的矩形区域内(图 3(b)),树高约 20 米。采用文献[26]的方法去除冠层点云后,剩余树桩和灌木点云数据,由 1802995 个离散点构成。从中划出点云密度较大的一个(30 米 × 31 米)矩形区域,有 261698 个采样点,记为 Γ ,如图 3(c)所示。然后,根据文献[26]的树干重建方法,从三维点云中识别出 68 棵树木和及其树根位置(图 3(d)中的蓝色点)。

从图 3(d) 可以看出, 尽管这是人工林的数据, 但是由于栽种施工误差等原因, 树木的位置只是大致成行布局, 并不形成规则的四边形网格。事实上, 这片毛白杨样地设计的是蜘蛛网形布局, 植株间距并不完全相同。由于树木死亡或者被砍伐等原因, 导致样地中有些树木缺失。缺失的数量和位置可以用本文第 3 节提出的 FMCP 算法方法进行检测和分析。本次实验中, 直线通过最少点数 k 取值为 4, 参数 ρ 与 ε 的取值同模拟实验, 区域 Γ 为已有离散点 p 形成的凸包(如图 3(d) 中虚线所示)。

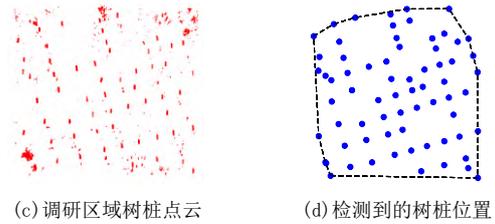
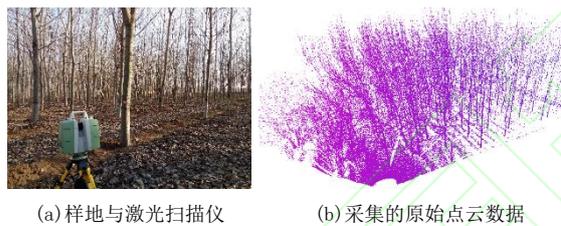


图 3 采集的数据与处理以及获取的树木位置
Fig.3 The Data Scanned in a Sampling Plot, Data Process and the Detected Tree Positions



采用 FMCP 和 Hough+ 算法的实验结果如图 4 所示, 其中 Hough+ 算法展示两个不同参数的实验效果 ($T=40$ 和 60)。图中实线条表示检测到的直线, 蓝色点为原有树的位置, 红色圆环点为算法识别出来的缺失树的位置。

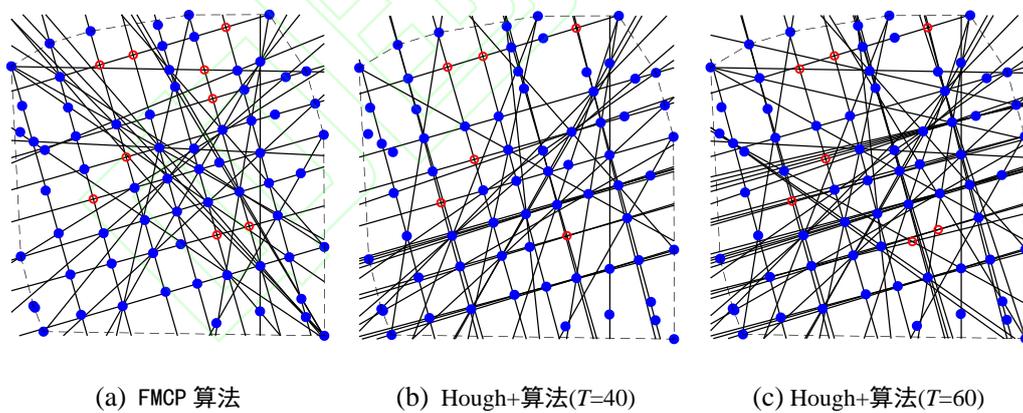


图 4 不同算法对于人工林缺失数据填补实验结果比较

Fig. 4 Comparison of Results Using Different Algorithms for Filling Missing Data in Plantations Point Cloud

使用 Hough 变换检测到的直线存在部分直线冗余, 即两条直线近似重合或相交形成的夹角过小, 而且有较多的漏检。这会影响到后续缺失数据填补效果。而使用 k -LDA 算

法检测出 37 条直线, 冗余直线极少且漏检直线少。

$T = 40$ 的 Hough+ 算法只检测到 6 棵缺失树的位置。缺失位置填补后, 其共线度从

原来的 0.1814 增大为 0.2152。 $T = 60$ 的 Hough+算法检测到 7 棵缺失树的位置,但这 7 棵树的位置只有部分处于人工林原有树位置构成的网格节点处。缺失位置填补后,位置点集共线度从原来的 0.2076 增大为 0.2459。而本文算法(FMCP)从该人工林数据中检测出 9 棵缺失树的位置,这 9 棵树的位置均大致处于人工林原有树位置所在的网格节点处。加入这 9 棵树后,该人工林的树位置之间的共线关系得到加强,其共线度从原来的 0.2193 增大为 0.2705。可见,本文提出的 FMCP 算法检测缺失数据位置的准确率明显高于 Hough+算法,且缺失位置填补后,点集共线度提升效果优于 Hough+算法。

5. 结论

本文针对具有共线特征的离散点集数据,定义了离散点集共线度的概念,提出了离散点集共线度的计算方法,并基于共线度最大化的优化目标结合 k -直线检测,提出一种二维缺失数据填补算法(FMCP)。实验结果表明,本文的缺失数据填补算法能有效地识别具有共线特征的缺失数据,在检测的直线质量和缺失数据填补效果方面优于基于 Hough 变换构建的方法。

本文算法主要适用于具有多重共线特

征的缺失位置检测和填补。对于非多重共线数据或者是曲线上的点的缺失分析未展开研究,因此,未来的一个有意义的研究方向是对缺失点位置的假设作出推广,将缺失点位置由 k -直线交点推广到 k -直线与曲线的交点,曲线与曲线交点等,使内蕴几何结构信息的缺失数据分析得到更多的关注。本文算法的另一个局限性在于近似共线直线的合并步骤的算法效率比较低,这是值得进一步研究的一个有工程价值的技术难点。

参考文献

- [1] Luo Yubo, Huang Hongyu, Tang Liyu, et al. Tree Height and Diameter Extraction with 3D Reconstruction in a Forest Based on TLS[J]. *Remote Sensing Technology and Application*, 2019, 34(2): 243-252 (骆钰波, 黄洪宇, 唐丽玉, 等. 基于地面激光雷达点云数据的森林树高、胸径自动提取与三维重建[J]. *遥感技术与应用*, 2019, 34(2): 243-252)
- [2] Roberts J, Tesfamichael S, Gebreslasie M, et al. Forest Structural Assessment Using Remote Sensing Technologies: An Overview of the Current State of the Art[J]. *Southern Hemisphere Forestry Journal*, 2007, 69(3): 183-203
- [3] Li W K, Guo Q H, Jakubowski M K, et al. A New Method for Segmenting Individual Trees from the Lidar Point Cloud[J]. *Photogrammetric Engineering & Remote Sensing*, 2012, 78(1): 75-84
程学院学报, 2017, 30(03): 37-42.)
- [4] Wang Shangshang, Wang Jun, Sun Si. Landscape

- Trees Counting by Remote Sensing Technology Based on UAV[J]. *Journal of Zhongkai University of Agriculture and Engineering*, 2017, 30(3): 37-42 (王上上, 王军, 孙思. 无人机遥感技术对园林树木的计数统计[J]. 仲恺农业工程学院学报, 2017, 30(3): 37-42)
- [5] Wu Tiandong, Qi Shu. Identification and Statistical Study on Tree Number of Eucalyptus Plantation Based on OpenCV[J]. *Forestry Science & Technology*, 2020, 45(5): 37-38 (吴天冬, 戚澍. 基于 OpenCV 的桉树人工林林木株数识别与统计研究[J]. 林业科技, 2020, 45(5): 37-38)
- [6] Cao Wei, Chen Dong, Shi Yufeng, et al. Progress and Prospect of LiDAR Point Clouds to 3D Tree Models[J]. *Geomatics and Information Science of Wuhan University*, 2021, 46(2): 203-220 (曹伟, 陈动, 史玉峰, 等. 激光雷达点云树木建模研究进展与展望[J]. 武汉大学学报·信息科学版, 2021, 46(2): 203-220)
- [7] Zhang X P, Li H J, Dai M R, et al. Data-Driven Synthetic Modeling of Trees[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2014, 20(9): 1214-1226
- [8] Liu Yawen, Zhang Ying, Chen Quan. Vehicle Point Cloud Data Enhancement Method Combined with Panoramic Image[J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(7): 1015-1020 (刘亚文, 张颖, 陈泉. 结合全景影像的车载街景点云数据增强方法[J]. 武汉大学学报·信息科学版, 2020, 45(7): 1015-1020)
- [9] Izzuddin R M, Pien C F, Jedol D, et al. Missing Value Imputation for PM₁₀ Concentration in Sabah Using Nearest Neighbour Method (NNM) and Expectation-Maximization (EM) Algorithm[J]. *Asian Journal of Atmospheric Environment*, 2020, 14(1): 62-72
- [10] Ozturk A. Accuracy Improvement in Air - Quality Forecasting Using Regressor Combination with Missing Data Imputation[J]. *Computational Intelligence*, 2021, 37(1): 226-252
- [11] Dzulkalnine M F, Sallehuddin R. Missing Data Imputation with Fuzzy Feature Selection for Diabetes Dataset[J]. *SN Applied Sciences*, 2019, 1(4): 362
- [12] Leng Yonglin, Zhang Qingchen, Lu Fuyu. Incomplete Big Data Imputation Algorithm Based on AP Clustering[J]. *Computer Engineering and Applications*, 2015, 51(10): 123-127 (冷泳林, 张清辰, 鲁富宇. 基于 AP 聚类的不完整大数据填充[J]. 计算机工程与应用, 2015, 51(10): 123-127)
- [13] Zhang Debing. *Study of Data Completion, Labeling and Retrieval Based on Machine Learning*[D]. Hangzhou: Zhejiang University, 2015 (张德兵. 基于机器学习的数据补全、标注和检索若干问题研究[D]. 杭州: 浙江大学, 2015)
- [14] Ye C, Wang H Z, Lu W B, et al. Effective Bayesian-Network-Based Missing Value Imputation Enhanced by Crowdsourcing[J]. *Knowledge-Based Systems*, 2020, 190: 105199
- [15] Miao Yongwei, Liu Jiazong, Chen Jiahui, et al. Structure-Preserving Shape Completion of 3D Point Clouds with Generative Adversarial Network[J]. *Scientia Sinica (Informationis)*, 2020, 50(5): 675-691 (缪永伟, 刘家宗, 陈佳慧, 等. 基于生成对抗网络的点云形状保结构补全[J]. 中

- 国科学: 信息科学, 2020, 50(5): 675-691)
- [16] Wang Yongbo, Wang Yunjia, She Wenwen, et al. A Linear Features-Constrained, Plücker Coordinates-Based, Closed-Form Registration Approach to Terrestrial LiDAR Point Clouds[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(9): 1376-1384 (王永波, 汪云甲, 余雯雯, 等. 直线特征约束下利用 Plücker 坐标描述的 LiDAR 点云无初值配准方法[J]. 武汉大学学报·信息科学版, 2018, 43(9): 1376-1384)
- [17] Lu X H, Liu Y H, Li K. Fast 3D Line Segment Detection from Unorganized Point Cloud[EB/OL]. 2019: arXiv: 1901.02532. <https://arxiv.org/abs/1901.02532>
- [18] Schnabel R, Wahl R, Klein R. Efficient RANSAC for Point-Cloud Shape Detection[J]. *Computer Graphics Forum*, 2007, 26(2): 214-226
- [19] Qi C R, Litany O, He K M, et al. Deep Hough voting for 3D object detection in point clouds[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South).: 9276-9285
- [20] Li Houlong, Ma Liming, Zhong Linwei. Lane Detection Based on Improved Hough Transform[J]. *Automobile Applied Technology*, 2021, 46(7): 16-18 (李后隆, 马骊溟, 钟林伟. 基于改进 Hough 变换的车道线检测[J]. 汽车实用技术, 2021, 46(7): 16-18)
- [21] Grompone von Gioi R, Jakubowicz J, Morel J M, et al. LSD: A Fast Line Segment Detector with a False Detection Control[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(4): 722-732
- [22] Li Hai, Zhang Xianmin, Chen Zhong. Automatic Corner Detection of Checkerboard Based on LSD[J]. *Optics and Precision Engineering*, 2015, 23(12): 3480-3489 (李海, 张宪民, 陈忠. 基于直线检测的棋盘格角点自动提取[J]. 光学精密工程, 2015, 23(12): 3480-3489)
- [23] Zhang Z C, Ma X. Lane Recognition Algorithm Using the Hough Transform Based on Complicated Conditions[J]. *Journal of Computer and Communications*, 2019, 07(11): 65-75
- [24] Yan Zhaojin, Meng Lina. Automatic Extraction Method of Road Sideline Based on Linear Feature[J]. *Engineering of Surveying and Mapping*, 2017, 26(3): 42-45 (闫兆进, 孟丽娜. 基于直线特征检测的道路边线自动提取方法[J]. 测绘工程, 2017, 26(3): 42-45)
- [25] Chen Baolin. *Optimization theory and algorithm*[M]. 2nd ed. 2nd ed. Beijing: Tsinghua University Press, 2005 (陈宝林. 最优化理论与算法[M]. 2版. 北京: 清华大学出版社, 2005)
- [26] Li H J, Zhang X P, Jaeger M, et al. Segmentation of Forest Terrain Laser Scan Data[C]//Proceedings of the 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry. Seoul, South Korea. 2010: 47-54

Counting of Plantation Trees Based on Line Detection of Point Cloud Data

FANG Hao¹, LI Hongjun¹

¹ College of Science, Beijing Forestry University, Beijing 100083, China

Abstract: Objectives: When performing plantation surveys using laser point cloud data, there are missing points in the scanned point cloud data due to the occlusion and self-occlusion of trees during laser scanning, the felling of trees and other reasons. So, the locations of the missing trees are inaccurate, and the forest survey results have large errors. The key to solving this problem is to realize the filling of the missing tree point cloud. Methods: This paper defines a concept named degree-of-collinearity, and constructs a method based on degree-of-collinearity combined with straight line detection to fill in missing data. Results: For the experimental results of simulated data, the average accuracy of the proposed algorithm is 97.28%; for the experimental results of plantation data, the proposed algorithm detects the location of 9 missing trees, and the degree-of-collinearity rises from 0.2193 to 0.2705. Conclusions: The experimental results show that this method can realize the optimal inference of missing location, strengthen the collinear relationship of filled data and can also be applied to count the missing trees in the artificial forest.

Key words: discrete point set; line detection; missing data completion; tree count; degree of collinearity

First author: FANG Hao, postgraduate, specializes in the analysis and processing of 3D point clouds. E - mail: 2505119182@qq.com

Corresponding author: LI Hongjun, PhD, Professor. E - mail: lihongjun69@bjfu.edu.cn

Foundation support: This research was supported by the 2020 Postgraduate Curriculum Construction Project of Beijing Forestry University (HXKC2005); The National Natural Science Foundation of China (61571046).

网络首发:

标题: 点云数据直线检测及其在人工林树木计数中的应用

作者: 方浩, 李红军

DOI: 10.13203/j.whugis20210407

收稿日期: 2021-07-29

引用格式:

方浩, 李红军. 点云数据直线检测及其在人工林树木计数中的应用[J]. 武汉大学学报·信息科学版, 2022, DOI: 10.13203/j.whugis20210407 (FANG Hao, LI Hongjun. Counting of Plantation Trees Based on Line Detection of Point Cloud Data[J]. *Geomatics and Information Science of Wuhan University*, 2022, DOI: 10.13203/j.whugis20210407)

网络首发文章内容和格式与正式出版会有细微差别, 请以正式出版文件为准!

您感兴趣的其他相关论文:

直线特征约束下利用 Plücker 坐标描述的 LiDAR 点云无初值配准方法

王永波, 汪云甲, 余雯雯, 韩新哲

武汉大学学报 信息科学版, 2018, 43(9): 1376-1384

<http://ch.whu.edu.cn/cn/article/doi/10.13203/j.whugis20160408>